SYNCHRONIZING LOCAL NETWORK

[0001] This is a Continuation of International Application No. PCT/FI02/00746 filed September 19, 2002, which designated the U.S. and was published under PCT Article 21(2) in English.

BACKGROUND OF THE INVENTION

Field of the Invention:
[0002] The present invention relates to telecommunications. In particular, the present invention relates to a novel and improved system for transmitting internal messages in a local network.

Description of the Related Art:
[0003] Fault tolerant local computer networks are used when high reliability is required. Typical examples are modern telecommunication networks where system downtime must be minimized in order to provide service of acceptable quality. A common way to implement fault tolerance is to use replicated elements in the networks. A network with replicated application processes is a network where each application process is simultaneously running in at least two separate computer units. Thus, if a computer unit running a given application process goes down for whatever reason, there's still at least one more computer unit running the same process.

[0004] Message synchronism is one possible way to implement a network with replicated elements. In other words, when a transmitting application process sends a message to a receiving application process, an identical copy of the message is sent to each replicated (or redundant) receiving process running in various computer units. Each replicating receiving process must receive an identical message and in identical order to maintain synchronization of messages. Message synchronism is sometimes also referred to as input synchronism.

[0005] A network with message synchronism can be implemented either by software or by hardware. When implemented by software, a special software layer is implemented on top of a standard messaging protocols or technologies such as Ethernet or IP (Internet Protocol, IP). However, software based messaging consumes a substantial amount of computing power and network bandwidth.

[0006] Another way to implement a network which takes care of the message synchronism is to use special proprietary hardware, typically a message bus. When a message is sent from one computer unit to another, it is replicated in the message bus hardware. The message bus acts as a serializing point and the message order is preserved.

[0007] A typical example of such a hardware based prior art implementation is illustrated in figure 1. The system comprises multiple sending computer units CPUs, each for running at least one sending application process for sending an internal message. Since the system in question implements fault tolerance with replicated elements, each message is sent to two or more recipients. The system further comprises multiple receiving computer units CPUr, each for running at least one receiving application process for receiving a sent internal message, at least two copies of each receiving application process residing in said receiving computer units.

[0008] The system further comprises one proprietary interface unit IF per one or more computer units for buffering and relaying messages sent to and from the corresponding computer unit. Each interface unit comprises a transmitting buffer TX for storing one or more message to be sent until processed by the internal bus, and a receiving buffer RX for storing one or more received messages until processed by the corresponding computer unit.

[0009] The system further comprises multiple internal links, such as e.g. Compact PCI –buses (Peripheral Component Interconnect, PCI), each for linking a computer unit to its corresponding interface unit.

[0010] The system further comprises a proprietary external message bus for receiving messages relayed by the interface units corresponding to the sending computer units, and for forwarding each received message to the interface units corresponding to the respective receiving computer units one received message at a time. The external message bus is linked with the interface units. The external message bus is a shared bus, i.e. only one sent message is forwarded by the bus at any given time, thus message synchronization is maintained.

[0011] However, there are significant disadvantages with the referenced prior art implementation. When more computing capacity is needed, more computer units are attached to the message bus. Thus a longer message bus is needed. Adding more computer units requires more messaging capacity between the computer units. The physical length of the message bus is limited because it is a shared bus. There is a trade-off between the length of the bus and the speed (bits per second). The higher the speed, the shorter the bus must be. The speed of the bus can be made higher by using more bits in parallel but the physical limits such as connectors and cabling will be reached very soon. Typically, the existing implementations use 8 or 16 bits in parallel.

[0012] Thus there is need for a solution overcoming the capacity limitations of the present message bus based implementations of a fault tolerant, high capacity synchronized networks.

[0013] Because the functionalities required from the system for providing message synchronism are in direct violation of present related specifications, commercially available solutions, e.g. Ethernet switches are not suitable.

This stems from the fact that the specification IEEE802.1D prohibits the switch sending a message back to the same computer unit that the message was received from. Yet such a functionality is required in order to provide message synchronism in all situations. Thus current implementations, such as the one disclosed in fig. 1, need to use proprietary components in order to achieve the desired functionalities.

[0014] Thus there is need for a solution making it possible to implement a system providing message synchronism using non-proprietary components either as-is or with only slight modifications.

SUMMARY OF THE INVENTION

[0015] The present invention concerns a system for transmitting internal messages in a local network while maintaining message synchronism. An internal message refers to a message sent from a computer unit in a local network to computer units in the same local network as opposed to traffic to/from external networks such as e.g. Internet. The system comprises multiple sending computer units, each for running at least one sending application process for sending an internal message. The system further comprises multiple receiving computer units, each for running at least one receiving application process for receiving a sent internal message, at least two copies of each receiving application process residing in said receiving computer units. The term receiving application process refers to application processes that for whatever reason are interested in receiving a sent message in identical order. In an embodiment of the present invention the receiving application processes are an application process and its replicated copies. The division of computer units into sending and receiving ones is a functional one, thus a computer unit can simultaneously be both sending and receiving messages. Correspondingly, an application process can simultaneously be both sending and receiving messages. The computer units are physically or logically separate from each other.

[0016] According to the invention, the system further comprises one interface unit per one or more computer units for buffering and relaying messages sent to and from the corresponding computer units.

[0017] Further according to the invention, the system comprises multiple external links, each for linking a computer unit to its corresponding interface unit. In an embodiment of the invention said external links are fast serial links.

[0018] Further according to the invention, the system comprises an internal interconnecting device for receiving messages relayed by the interface units corresponding to the sending computer units, and for forwarding each received message to the interface units corresponding to the respective receiving computer units one received message at a time thus maintaining message synchronism. The interconnecting device is internally coupled with the interface units.

[0019] Thus also messages sent and received by application processes running in the same computer unit are routed via the interconnecting device. Since the long and thus relatively slow external message bus of prior art is replaced with a short and extremely fast internal interconnecting device, the overall messaging capacity of the system increases while message synchronization is maintained. Yet, because of external links used between computer units and interface units, distance between the computer units can still be relatively long, i.e. tens of meters.

[0020] Another embodiment of the invention concerns a system for transmitting internal messages in a local network while maintaining message synchronism. The system comprises multiple sending computer units, each for running at least one sending application process for sending an internal message. The system further comprises multiple receiving computer units, each for running at least one receiving application process for receiving a

sent internal message, at least two copies of each receiving application process residing in said receiving computer units. The computer units are physically or logically separate from each other.

[0021] According to this embodiment of the invention, the system comprises multiple multiplexer units, each for collecting messages from and distributing messages to one or more sending computer units.

[0022] Further according to this embodiment of the invention, the system further comprises one interface unit per one or more multiplexer units for buffering and relaying messages sent to and from the corresponding multiplexer units. In an embodiment of the invention there may be multiple layers of multiplexer units, thus a multiplexer unit may be connected to the interface unit via another multiplexer unit. By doing so the number of computer units linked to the internal interconnecting device using one interface unit can be increased. At the same time the amount of cabling may be reduced.

[0023] Further according to this embodiment of the invention, the system comprises an internal interconnecting device for receiving messages relayed by the interface units corresponding to the sending computer units, and for forwarding each received message to the interface units corresponding to the respective receiving computer units one received message at a time thus maintaining message synchronism. The interconnecting device is internally coupled with the interface units. Because of multiplexers used between computer units and interface units, cabling is reduced significantly.

[0024] In an embodiment of the invention each interface unit further comprises a transmitting buffer for storing one or more messages to be sent until processed by the interconnecting device, and a receiving buffer for storing one or more received messages until processed by the corresponding computer or multiplexer unit.

[0025]   In an embodiment of the invention messages are sent as multicasts by the sending application process. Multicasting can be done e.g. by using group addressing. Group addressing refers to a known addressing method where a group of computer units are assigned a common group address. A message to be sent to computer units of a given group is addressed according to corresponding group address. With two recipients group addressing is commonly referred to as pair addressing.

[0026]   In an embodiment of the invention the interconnecting device is an internal bus.

[0027]   In an embodiment of the invention the interconnecting device is a crossbar.

[0028]   In an embodiment of the invention the interconnecting device, the interface units coupled to it and the multiplexer units are implemented by modifying a standard LAN switch such as an Ethernet switch. Such modifications comprise modifications making it possible to multicast messages back to the same computer unit that the message was received from, as well as modifications required to make the interconnecting device act as a serializing point.

[0029]   The invention makes it possible to overcome the capacity limitations of the present external message bus based proprietary implementations of a fault tolerant, high capacity synchronized networks. Significantly higher messaging capacity can be achieved than with existing implementations while at the same time maintaining the message synchronization required by the fault tolerant, high capacity synchronized networks. Additionally, the invention makes it possible to implement a system providing message synchronism using non-proprietary components either as-is or with only slight modifications. The system can be implemented e.g. by only slightly modifying already existing LAN switches.

BRIEF DESCRIPTION OF THE DRAWINGS

[0030]   The accompanying drawings, which are included to provide a further understanding of the invention and constitute a part of this specification, illustrate embodiments of the invention and together with the description help to explain the principles of the invention. In the drawings:

[0031]   **Fig 1** is a block diagram illustrating a prior art system,

[0032]   **Fig 2** is a block diagram illustrating a system according to one embodiment of the present invention, and

[0033]   **Fig 3** is a block diagram illustrating a system according to another embodiment of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0034]   Reference will now be made in detail to the embodiments of the present invention, examples of which are illustrated in the accompanying drawings.

[0035]   Figure 2 illustrates a system for transmitting internal messages in a local network while maintaining message synchronism. The system comprises multiple sending computer units CPUs (only one of which is disclosed in Figure 2), each for running at least one sending application process for sending an internal message. The system further comprises multiple receiving computer units CPUr, each for running at least one receiving application process for receiving a sent internal message, at least two copies of each receiving application process residing in said receiving computer units. Application processes are elements performing information processing for a given application.

[0036]   The system illustrated in Figure 2 further comprises one interface unit IF per one or more computer units for buffering and relaying messages

sent to and from the corresponding computer units. In the embodiment of the invention illustrated in Figure 2 messages are sent as multicasts by the sending application process. Each interface unit IF further comprises a transmitting buffer TX for storing one or more messages to be sent until processed by the interconnecting device, and a receiving buffer RX for storing one or more received messages until processed by the corresponding computer unit. Preferably the buffers are FIFO (First In – First Out) buffers.

[0037] The system illustrated in Figure 2 further comprises multiple external links SrL, each for linking a computer unit to its corresponding interface unit. In the embodiment of the invention illustrated in Figure 2 said external links are fast serial links.

[0038] The system illustrated in Figure 2 further comprises an internal interconnecting device IxD for receiving messages relayed by the interface units corresponding to the sending computer units, and for forwarding each received message to the interface units corresponding to the respective receiving computer units one received message at a time thus maintaining message synchronism. The interconnecting device is internally coupled with the interface units. In the embodiment of the invention illustrated in Figure 2 said interconnecting device is an internal message bus. However, the interconnecting device can also be implemented e.g. as a crossbar or as some other internal bus arrangement known to a person skilled in the art. Messages sent and received by application processes running in the same computer unit are also routed via the internal bus.

[0039] Figure 3 illustrates a system for transmitting internal messages in a local network while maintaining message synchronism. The system comprises multiple sending computer units CPUs, each for running at least one sending application process for sending an internal message. The system further comprises multiple receiving computer units CPUr, each for running

at least one receiving application process for receiving a sent internal message, at least two copies of each receiving application process residing in said receiving computer units. Application processes are elements performing information processing for a given application.

[0040] The system illustrated in Figure 3 further comprises multiple multiplexer units MUX, each for collecting messages from and distributing messages to one or more sending computer units. The multiplexers collect traffic and messages from one or more computer units and deliver them to the internal interconnecting device via the interface units. The multiplexers are used to reduce cabling and the number of interfaces or ports in the interface units and the interconnecting device. The multiplexers may be implemented so that they pass all the traffic from computer units to interface units. Additionally the multiplexers can be implemented so that unicast messages sent to the computer units residing in the area of the same multiplexer are directed to the correct computer units. However in case of multicast messages, if any of the addressed computer units or application processes is not residing in the area of said multiplexer, the message is directed to the right addresses via the interface units and the interconnecting device. Otherwise the message synchronism can not be maintained. In the embodiment of the invention illustrated in Figure 3 there is one layer of multiplexer units. However, the invention can also be implemented with multiple layers of multiplexer units.

[0041] The system illustrated in Figure 3 further comprises one interface unit IF per one or more multiplexer units for buffering and relaying messages sent to and from the corresponding multiplexer units. In the embodiment of the invention illustrated in Figure 3 messages are sent as multicasts by the sending application process. Each interface unit IF further comprises a transmitting buffer TX for storing one or more messages to be sent until processed by the interconnecting device, and a receiving buffer RX for stor-

ing one or more received messages until processed by the corresponding multiplexer unit. Preferably the buffers are FIFO (First In – First Out) buffers.

[0042] The system illustrated in Figure 3 further comprises an internal interconnecting device IxD for receiving messages relayed by the interface units corresponding to the sending computer units, and for forwarding each received message to the interface units corresponding to the respective receiving computer units one received message at a time thus maintaining message synchronism. The interconnecting device is internally coupled with the interface units. In the embodiment of the invention illustrated in Figure 3 said interconnecting device is an internal message bus. However, the interconnecting device can also be implemented e.g. as a crossbar or as some other internal bus arrangement known to a person skilled in the art. Messages sent and received by application processes running in the same computer unit are also routed via the internal bus.

[0043] It is obvious to a person skilled in the art that with the advancement of technology, the basic idea of the invention may be implemented in various ways. The invention and its embodiments are thus not limited to the examples described above, instead they may vary within the scope of the claims.